Psychology of Sport and Exercise 16 (2015) 150–159

Contents lists available at ScienceDirect

Psychology of Sport and Exercise

journal homepage: www.elsevier.com/locate/psychsport

The TRAMS: The Team-Referent Attributions Measure in Sport

Pete Coffee [a,*], Iain Greenlees [b], Mark S. Allen [c]

[a] School of Sport, University of Stirling, Stirling, FK9 4LA, UK
[b] Department of Sport and Exercise Sciences, University of Chichester, Bishop Otter Campus, College Lane, Chichester, PO19 6PE, UK
[c] School of Psychology, Faculty of Social Sciences, University of Wollongong, Northfields Avenue, Wollongong, New South Wales, 2522, Australia

ARTICLE INFO

Article history:
Received 29 May 2014
Received in revised form
24 October 2014
Accepted 24 October 2014
Available online 3 November 2014

Keywords:
Confirmatory factor analysis
Construct validity
Group dynamics
Multilevel analyses
Sport psychology

ABSTRACT

Objectives: To provide initial evidence for the construct, concurrent, and predictive validity of the Team-Referent Attributions Measure in Sport (the TRAMS).
Design: Cross-sectional in Studies 1 and 2, and multiple time points in Study 3.
Method: Study 1 required participants ($N = 500$) to complete the TRAMS for their “least successful” and “most successful” performances in the preceding three months. In Study 2, after performance, participants ($N = 515$) completed the TRAMS and the Causal Dimension Scale for Teams (CDS-T; Greenlees et al., 2005). Study 3 required participants ($N = 165$) to complete a measure of pre-competition collective-efficacy prior to performance (Day 1, Time 1), the TRAMS following performance (Day 1, Time 2), and a measure of subsequent collective-efficacy prior to subsequent performance (Day 7–9, Time 3).
Results: Study 1 supported the factor structure of the TRAMS across least successful and most successful conditions. Study 2 provided further support for the factor structure of the TRAMS, together with evidence of concurrent validity with subscales of the CDS-T. Study 3 revealed, following team defeat, interactions between controllability and generalisability dimensions: Controllability had a significant effect upon subsequent collective-efficacy when causes of team defeat were also perceived to generalise across situations and/or across teams. Following team victory, stable attributions were positively associated with subsequent collective-efficacy.
Conclusions: This article provides initial evidence for the validity of the TRAMS and demonstrates for team-referent attributions the theoretical advantages of examining a broader conceptualisation of generalisability attributions and interactive effects of attributions.

Team-referent attributions refer to team members' individual explanations for collective outcomes. The attributions made for group outcomes are proposed to have an important role in the affective, cognitive, physiological and behavioural responses of group members (Allen, Coffee, & Greenlees, 2012). For example, the explanations we form for group outcomes influence emotions (e.g., pride or shame), efficacy for future performances (i.e., confidence to perform in the future), hormones (e.g., elevation or reduction in cortisol and testosterone levels), and subsequent behaviour (e.g., increased or reduced involvement).

A central premise within attribution research is that there is a dimensional structure underpinning the reasons people give for their failures and successes. The Attribution Theory of Achievement Motivation (ATAM; e.g., Weiner, 1985) considers three primary attribution dimensions: locus of causality, stability, and controllability. Locus of causality refers to the extent to which causes are seen as either residing within or outside an individual. Stability refers to the extent to which causes are seen as either stable or variable. Controllability refers to the extent to which causes are seen as regulated by individuals/teams or something over which control cannot be exerted. For example, an attribution to “poor genetics” is usually (but not always) categorised as internal, stable, and uncontrollable.

An important advancement in the measurement of *team-referent attributions*, based upon the ATAM model, was the development of the Causal Dimension Scale for Teams (CDS-T; Greenlees, Lane, Thelwell, Holder, & Hobson, 2005). This self-assessment, situation-specific questionnaire, derived from the Revised Causal Dimension Scale (CDS-II; McAuley, Duncan, & Russell, 1992), requires team members to identify the main cause of a team outcome and then to rate that cause along a series of bipolar items that correspond to four attribution dimensions: locus of causality, stability, team control (control by the team), and external control

* Corresponding author. Tel.: +44 (0)1786 466253; fax: +44 (0)1786 466477.
E-mail addresses: peter.coffee@stir.ac.uk (P. Coffee), i.greenlees@chi.ac.uk (I. Greenlees), mark_allen@uow.edu.au (M.S. Allen).

http://dx.doi.org/10.1016/j.psychsport.2014.10.009
1469-0292/

(control by others). To develop the CDS-T, Greenlees and colleagues made two major amendments to the CDS-II: First, the authors reworded each item to reflect a team rather than a self-referent attribution; the word "you" was replaced with "your team" throughout the revised questionnaire. Second, the authors generated four new items to increase the item pool for each attribution dimension from three items to four items. Greenlees et al. reported good overall fit for the four-factor CDS-T ($\chi^2(98) = 210.21$; RMSEA = .05, $p > .05$; NNFI = .95; CFI = .96) with coefficient alpha reliabilities ranging from .74 to .82. There were, however, some concerns about low factor loadings for items with six factor loadings below .70 and one factor loading below .50.

The CDS-T has subsequently been used to explore the correlates of team-referent attributions in sport settings (Allen, Jones, & Sheffield, 2009; Chow & Feltz, 2008; Dithurbide, Sullivan, & Chow, 2009; Greenlees et al., 2007; Shapcott, Carron, Greenlees, & El Hakim, 2010). For example, Greenlees et al. demonstrated that attributions following success were more likely to be perceived as internal, stable, and controllable, and Dithurbide et al. reported evidence that collective efficacy was higher when causes of prior performance were considered less stable. Allen and colleagues explored interactive effects of attributions on collective-efficacy and reported, following team defeat, an interaction for external control and stability, and, following team victory, an interaction for team control and stability. Across conditions, the nature of the interactions was the same: If causes were perceived as stable, higher levels of control (team control—following team defeat—and external control—following team victory) were positively associated with subsequent collective-efficacy.

Despite the progress in the team-referent attribution literature—afforded through the development of the CDS-T—theoretical advancements are limited by the lack of an alternative validated measure to the CDS-T for assessing situation-specific team-referent attributions in sport. This reasoning is underpinned by the following two key points: First, the CDS-T suffers from the same conceptual and measurement issues inherent in its parent measure, the CDS-II. In reference to the CDS-II, it has been noted that the assessment of personal (team, for the CDS-T) and external control is not congruent with the ATAM framework and that respondents have considerable problems interpreting scale anchors (Biddle & Hanrahan, 1998; Biddle, Hanrahan, & Sellars, 2001). Typically, with research using the CDS-II, the conceptual modification of controllability results in very high correlations between locus of causality and personal control (e.g., Crocker, Eklund, & Graham, 2002). Similarly, very high and significant correlations between locus of causality and team control have been noted in research using the CDS-T ($rs = .69–.79$, $ps < .01$, Allen et al., 2009; $r = .84$, $p < .01$, Dithurbide et al., 2009), suggesting cause for concern regarding the discriminant validity of the subscales. Second, empirical evidence from research into self-referent attributions (Coffee & Rees, 2008a, 2008b, 2009; Crocker et al., 2014) provides support for a broader conceptual approach to assessing attributions in sport (Allen et al., 2012; Rees, Ingledew, & Hardy, 2005).

Rees et al. (2005) proposed that attribution research in sport should focus upon the main effects of controllability, together with the interactive effects of controllability and generalisability dimensions (stability, globality, and universality). This proposal is underpinned by three key points. First, that controllability is a key dimension upon which attention should be focused. Second, that attribution research would benefit from examining a broader conceptualisation of generalisability dimensions; that is, in addition to stability, examining the globality and universality of causes. As noted before, stability refers to the extent to which causes are seen as either stable or variable. Globality refers to the extent to which causes are seen to affect a wide range of situations or a narrow range of situations, and universality refers to the extent to which causes are seen as common to all people/teams or unique to individuals/teams. Third, that to model generalisability implies the need to move beyond main effects and consider interactive effects of attribution dimensions. (The reader is referred to Coffee & Rees, 2008a, and Rees et al., 2005, for a more elaborate discussion of these proposals.)

To permit investigation of these proposals, Coffee and Rees (2008a) developed a measure of Controllability, Stability, Globality, and Universality attributions (the CSGU). The CSGU is a 16-item self-referent, situation-specific self-report questionnaire that assesses controllability and the three generalisability dimensions of stability, globality, and universality. The authors reported a good fit for the four-factor structure (least successful condition: RMSEA = .04, SRMR = .04, NNFI = .98; CFI = .98; most successful condition: RMSEA = .04, SRMR = .05, NNFI = .97; CFI = .98), together with coefficient alpha reliabilities ranging from .81 to .91. In line with proposals by Rees et al. (2005), the CSGU has been used to explore the main and interactive effects of attribution dimensions upon emotional and cognitive consequences of attributions (Coffee & Rees, 2008a, 2008b, 2009; Crocker et al., 2014). For example, Crocker and colleagues reported significant relationships between guilt and attributions of controllability, stability and globality, and between physical self-concept and shame and attributions of controllability and globality. Coffee and Rees have reported, following less successful performances, interactive effects for controllability and generalisability dimensions, and, following more successful performances, main effects for generalisability dimensions upon self-efficacy. The interactive effects, following less successful performances, demonstrated that controllability was positively associated with subsequent self-efficacy when causes were perceived to generalise across time (stability; Coffee & Rees, 2008a, 2009) or situations (globality; Coffee & Rees, 2008b).

The purpose of the current article is to provide initial evidence for the construct, concurrent, and predictive validity of a new, four-factor (controllability, stability, globality, and universality) measure of team-referent attributions: the Team-Referent Attributions Measure in Sport (the TRAMS). The TRAMS assesses individual group member perceptions of the causes of group outcomes. In Study 1, we examined the construct validity of the TRAMS across least successful and most successful conditions using confirmatory factor analysis (CFA). In Study 2, following team defeat and team victory, we again tested the factor structure of the TRAMS through CFA, together with examining evidence for concurrent validity by exploring correlations between the TRAMS dimensions and dimensions of the CDS-T. We hypothesized that high correlations would emerge between the TRAMS controllability subscale and the CDS-T team control subscale, and between the TRAMS stability subscale and the CDS-T stability subscale. In Study 3 we explored the predictive validity of the TRAMS and examined main and interactive effects of team-referent attributions on collective-efficacy following team defeat and team victory.

Study 1

Method

Participants

Participants were 500 (121 female; M_{age} 19.88 ± 2.04 years) sport, exercise, and health science students at three universities in the UK who competed in sport. The sample was predominantly White (85.40%; 28 participants did not report ethnicity). Participants had competed for a mean of 9.99 (SD = 4.18) years in their main sport. Participants self-selected their level of competition/performance from the descriptors club ($n = 263$), county ($n = 108$), regional ($n = 69$), national ($n = 39$), and international ($n = 17$) level (four participants did not report their level of performance). The principal sports of participants included soccer ($n = 246$), rugby ($n = 54$), cricket ($n = 39$), basketball ($n = 25$), and hockey ($n = 25$).

Procedure

Ethical approval was obtained from a university ethics review committee and participants provided informed consent. Participants completed the questionnaire before or after a lecture, and participation was voluntary with no course credits or financial incentives offered. Team sport athletes were asked to participate and the questionnaire took approximately 10 min to complete. Participants were asked to remember their perceived least successful team performance within the past three months (condition 1) before answering the following question, "To what extent was this performance successful in comparison to your team's general performance level?" Response options ranged from 1 (*not at all*) to 5 (*completely*). With this performance in mind, an open ended statement required participants to write down the single most important reason to explain their team's performance. In relation to this reason, participants completed the TRAMS. Immediately after, participants repeated this procedure for their team's most successful performance within the past three months (condition 2).[1]

Measures

A 16-item measure of team-referent attributions was used; the TRAMS. The TRAMS was based on the CSGU developed by Coffee and Rees (2008a) with a single major amendment. Where necessary, items were reworded to reflect a team-referent rather than a self-referent attribution. Specifically, the word "you" was replaced with "your team" and "athletes" was replaced with "team" throughout the revised questionnaire.[2] Further, the singular "encounter" was replaced with the plural "encounters" in the item "relates to a number of different situations your team encounters". As such, the TRAMS assessed the four subscales (four items per subscale) of controllability, stability, globality, and universality. Items were prefixed with the question, "In general, to what extent is your reason something that…" Examples of items are as follows: "your team could control in the future" (controllability), "remains stable across time" (stability), "affects a wide variety of outcomes for your team" (globality), and "is a common cause of performance for other teams" (universality). Participants' responses were recorded on a 1 (*not at all*) to 5 (*completely*) Likert scale, with higher values representing items that were more controllable, stable (except for the reverse scored item, "fluctuates across performances"), global, and universal.

Analyses

The data were screened for missing values and indices of non-normality. The factor structure of the TRAMS was tested using CFA with maximum likelihood estimation (Jöreskog & Sörbom, 1996).[3] Data analyses using LISREL 9.1 (Jöreskog & Sörbom, 2012) were conducted separately for the two conditions. The sequential model testing approach, as recommended by Jöreskog (1993) was employed and involved three stages. First, tests of separate single-factor models corresponding to individual subscales were performed, the purpose of which was to assess the convergent validity of the items making up each subscale. Overall fit indices of each model were considered along with the completely standardised factor loadings (loadings with values for z above 1.96 were considered significant), the standardised residuals (values above 2 and below −2 were considered large), and the modification indices for the covariances between measurement errors (values above 7 were considered large; Jöreskog & Sörbom, 1996). Second, tests of two-factor models were undertaken by combining each pair of attribution subscales. The purpose of this stage was to identify ambiguous items and investigate the discriminant validity of the factors. Large modification indices suggested that improvements in fit could be expected if items were freed to cross-load on another factor, and a confidence interval (95% CI; ±1.96 standard errors) including 1.0 suggested that the factors were perfectly correlated and therefore lacked discriminant validity. All factors were then included in a full four-factor model. The four-factor model was also tested for factorial invariance across conditions. As data were collected on one sample, a within-subject design was employed, testing invariance in one CFA by allowing corresponding factors, items, and error variances to covary (see, e.g., Raykov, 2006). CFA models fit to the data included configural, measurement, and structural factorial invariance models (see Byrne, 2006).

Following evidence of non-normality (multivariate skewness, $z = 24.00$, $p < .01$; multivariate kurtosis $= 18.32$, $p < .01$), the goodness of fit of all models was tested using the Satorra–Bentler chi-square statistic (SB χ^2), together with the Root Mean Square Error of Approximation (RMSEA) and its associated p-value (for RMSEA $< .05$), the Standardised Root Mean Square Residual (SRMR), the Comparative Fit Index (CFI), and the Non-Normed Fit Index (NNFI). These fit indices included measures from three different classes (absolute fit, absolute fit with penalty function, and incremental/comparative fit) (Hu & Bentler, 1999; Jöreskog, 1993). The SB χ^2 statistic was used as a subjective index of fit (Jöreskog & Sörbom, 1996). The recommendations for fit of Hu and Bentler are values for SRMR close to .08, RMSEA close to .06, and CFI and NNFI close to .95.[4] To compare fit across configural, measurement, and structural factorial invariance models, the difference in CFI and the Satorra–Bentler scaled chi-square different test (Satorra & Bentler, 2001) were examined.[5] We also assessed the coefficient alpha reliabilities and composite reliabilities of each of the factors in the final model. Composite reliability draws on the standardised loadings and measurement errors, with values above .70 indicating acceptable composite reliability (Shook, Ketchen, Hult, & Kacmar, 2004).[6] An alpha level of .05 was used for all tests.

Results

Fourteen participants were removed from analyses through listwise deletion for missing values for TRAMS items. Responses to the item "To what extent was this performance successful in comparison to your team's general performance level" revealed a

[1] A similar procedure was adopted in the development of the CSGU (see, Coffee & Rees, 2008a).

[2] A similar procedure was employed in the development of the CDS-T based upon the CDS-II (see Greenlees et al., 2005).

[3] The data in Study 1 were not nested in teams. The factor structure of the attributions measure was assessed for participants' perceived "least successful" and "most successful" team performances in the preceding three months. As such, participants could respond with a different team performance in mind.

[4] Browne and Cudeck (1993) suggested that values for RMSEA up to .08 indicate a reasonable error of approximation, but models with values greater than .10 would be unacceptable.

[5] There is an increasing tendency (see Byrne, 2006) to compare the fit of models based on two alternative criteria to $\Delta\chi^2$: (a) the value of ΔCFI between models is negligible (Cheung & Rensvold, 2002, suggested that a value of ΔCFI smaller than or equal to −.01 indicates that the null hypothesis of invariance should not be rejected), and (b) the overall model exhibits an adequate fit to the data.

[6] Composite reliability ρ_c is defined as (adapted from Fornell & Larcker, 1981): $\rho_c = (\sum L_i)^2 / \left((\sum L_i)^2 + \sum \mathrm{Var}(E_i)\right)$ where L_i is the standardised factor loadings for that factor, and $\mathrm{Var}(E_i)$ is the error variance associated with the individual indicator variables (items).

Table 1
Completely standardised solution and fit statistics for the full four-factor model in the least successful condition.

Items	Measurement error variances	Factor C	Factor S	Factor G	Factor U
		Item-factor loadings			
your team could control in the future	.70	.68			
in the future, your team could exert control over	.35	.80			
in the future, your team could change at will	.82	.60			
your team could regulate in the future	.40	.79			
remains stable across time	.61		.68		
you feel remains constant over time	.38		.81		
stays consistent across time	.50		.74		
relates to a number of different situations your team encounters	.57			.58	
affects a wide variety of outcomes for your team	.55			.62	
influences the outcomes of new situations your team face	.47			.64	
influences all situations your team encounters	.56			.65	
is a common cause of performance for other teams	.44				.74
is a cause of performance that other teams relate to	.44				.69
can be used to explain the performances of other teams	.54				.72
is a cause of performance for other teams as well	.41				.77

Factor	$M \pm SD$	ρ_c	α	Factor-factor correlations		
Controllability (C)	3.54 ± .85	.78	.80			
Stability (S)	2.47 ± .88	.77	.78	.21*		
Globality (G)	3.42 ± .69	.74	.72	.43*	.30*	
Universality (U)	3.48 ± .80	.82	.82	.37*	.17*	.59*

Full four-factor least successful model	SB χ^2	d.f.	p(SB χ^2)	RMSEA	RMSEA (p)	SRMR	CFI	NNFI
	162.79	84	< .01	.04	.83	.04	.98	.97

Note. $n = 486$. ρ_c = Composite reliability. α = Coefficient alpha. SB χ^2 = Satorra Bentler χ^2. RMSEA = Root Mean Square Error of Approximation. SRMR = Standardised Root Mean Square Residual. CFI = Comparative Fit Index. NNFI = Non-Normed Fit Index.
*$p < .01$.

significant difference ($t(482) = 29.29$, $p < .01$) between participants' perceptions of their team's least successful ($M = 1.90 \pm 1.05$) and most successful ($M = 3.65 \pm .80$) performances.[7]

Least successful condition

At the single-factor stage, the majority of chi-square statistics for model fits were non-significant (chi-square for universality was significant), RMSEA values ranged from <.01 to .08 (all were non-significant), SRMR values ranged from <.01 to .02, CFI values were .99 and 1.00, and NNFI values ranged from .98 to 1.01. Factor loadings were all significant except for the factor loading of .02 ($t = .38$, $p > .05$) for the stability item "fluctuates across performances." The item was retained at this stage and further explored at the two-factor stage. (Detailed information of the fit statistics at the single-factor stage is provided online in Supplementary Table 1.)

At the two-factor stage, the stability item "fluctuates across performances" exhibited high modification indices with all other factors (modification indices for lambda-*X* were 30.33, 35.21 and 22.77 for controllability, globality, and universality, respectively). Due to the low and non-significant loading of the item on its hypothesized factor and the high cross-loadings to all other factors, the item was removed. All two-factor models were good: RMSEA values ranged from .02 to .06 (all were non-significant), SRMR values from .03 to .05, CFI from .98 to 1.00, and NNFI from .97 to 1.00. The 95% confidence interval around two-factor intercorrelations ranged from .17 to .59. (Detailed information of the fit statistics for the final two-factor models, not including "fluctuates across performances", is provided online in Supplementary Table 2.)

[7] Three further participants were removed from this analysis due to missing values on the perceptions of success measure.

At the full four-factor model stage, although the chi-square statistic was significant (SB $\chi^2(84) = 162.79$, $p < .01$), the RMSEA was low (.04), with a non-significant test for close fit, the SRMR was low (.04), and the CFI (.98) and NNFI (.97) were high. These values are indicative of good fit (Hu & Bentler, 1999). Coefficient alpha reliabilities for the four subscales ranged from .72 to .80 and composite reliabilities ranged from .74 to .82. The completely standardised solution for the full four-factor model is presented in Table 1.

Most successful condition

The 15-item four-factor model identified in the least successful condition was tested for fit using data from the most successful condition. The completely standardised solution for the full four-factor model is presented in Table 2. Although the chi-square statistic was significant (SB $\chi^2(84) = 157.76$, $p < .01$), the RMSEA was low (.04), with a non-significant test for close fit, the SRMR was low (.05), and the CFI (.97) and NNFI (.97) were high. These values are indicative of good fit (Hu & Bentler, 1999). Coefficient alpha reliabilities ranged from .63 to .77 and composite reliabilities from .70 to .79.

Factorial invariance

First, an eight-factor model of covariance structures allowing corresponding factors, items, and error variances to covary was fitted to the data. This model (baseline model) imposed no equality constraints on parameter estimates across conditions. The model fit provided evidence for configural factorial invariance (SB $\chi^2(374) = 600.15$, $p < .01$; RMSEA = .04, $p = 1.00$; SRMR = .08; CFI = .97; NNFI = .97). The second model tested for measurement factorial invariance, constraining corresponding factor loadings to be equal across conditions. Although the SB scaled χ^2 difference test (SB $\chi^2(15) = 42.01$, $p < .01$) suggested a significant difference between the models, there was no change in the value for CFI and an

Table 2
Completely standardised solution and fit statistics for the full four-factor model in the most successful condition.

Items	Measurement error variances	Factor: C	S	G	U
		Item-factor loadings			
your team could control in the future	.53	.65			
in the future, your team could exert control over	.49	.61			
in the future, your team could change at will	.87	.47			
your team could regulate in the future	.43	.66			
remains stable across time	.47		.66		
you feel remains constant over time	.43		.76		
stays consistent across time	.55		.67		
relates to a number of different situations your team encounters	.54			.51	
affects a wide variety of outcomes for your team	.58			.47	
influences the outcomes of new situations your team face	.48			.51	
influences all situations your team encounters	.40			.68	
is a common cause of performance for other teams	.59				.62
is a cause of performance that other teams relate to	.51				.57
can be used to explain the performance of other teams	.53				.69
is a cause of performance for other teams as well	.29				.81

Factor	*M* ± *SD*	ρ_c	α	Factor-factor correlations: C	S	G
Controllability (C)	3.48 ± .68	.71	.68			
Stability (S)	3.05 ± .79	.75	.74	.33*		
Globality (G)	3.58 ± .58	.70	.63	.60*	.29*	
Universality (U)	3.47 ± .73	.79	.77	.36*	.21*	.61*

Full four-factor most successful model	SB χ^2	d.f.	p(SB χ^2)	RMSEA	RMSEA (p)	SRMR	CFI	NNFI
	157.76	84	< .01	.04	.88	.05	.97	.97

Note. $n = 486$. ρ_c = Composite reliability. α = Coefficient alpha. SB χ^2 = Satorra Bentler χ^2. RMSEA = Root Mean Square Error of Approximation. SRMR = Standardised Root Mean Square Residual. CFI = Comparative Fit Index. NNFI = Non-Normed Fit Index.
*$p < .01$.

excellent fit was observed for the model to the data (RMSEA = .04, p = 1.00; SRMR = .09; NNFI = .97). The third model tested for structural factorial invariance, constraining corresponding factor loadings and factor covariances to be equal across conditions. Although the SB scaled χ^2 difference test (SB $\chi^2(21) = 58.76$, $p < .01$) suggested a significant difference between the models, there was no change in the value for CFI and an excellent fit was observed for the model to the data (RMSEA = .04, p = 1.00; SRMR = .09; NNFI = .97).

Discussion

Following removal of the reverse scored stability item, "fluctuates across performances", the 15-item four-factor structure of the TRAMS was confirmed across least successful and most successful conditions. Across conditions, all factor loadings were significant. The item "fluctuates across performances" was removed due to a low and non-significant loading on its hypothesized factor and high cross-loadings to all other factors. The item was the only reverse scored item in the proposed 16-item TRAMS. Across conditions, factorial invariance analyses provided evidence of configural, measurement, and structural factorial invariance. In summary, the results of Study 1 provided initial support for the construct validity of a 15-item TRAMS (a copy of the final instrument is provided online in the Supplementary Material).

Study 2

Method

Participants

Participants were 515 (123 female; M_{age} 22.17 ± 4.70 years) competitive athletes. The sample was predominantly White (86.41%). Participants had competed for a mean of 9.81 (SD = 6.17) years in their main sport. Participants competed at club (n = 346), county (n = 71), regional (n = 52), national (n = 35), or international (n = 8) level (three participants did not report their level of performance).

Procedure

Ethical approval was granted by a university ethics committee and participants provided informed consent. Sampling was opportunistic with clubs informed about the study and with participants recruited at the site of competitions with the aid of an information sheet. Data were collected up to one hour after performance (e.g., a soccer match or a netball match) to give participants a chance to physically recover from competition. Participants were asked, "Would you consider your team's performance a failure or success?" with binary response options of "failure" and "success". An open-ended statement required participants to write down the single most important reason for their team's performance. In relation to this reason, participants completed two measures of attributions.

Measures

The 15-item four-factor TRAMS developed in Study 1 and the CDS-T (Greenlees et al., 2005) were used. The CDS-T comprises 16 items assessing the four subscales (four items per subscale) of locus of causality, stability, team control, and external control. Examples of items are as follows: "caused by an aspect of your team—caused by an aspect of the situation" (locus of causality), "permanent—temporary" (stability), "your team can control— your team cannot control" (team control), and "people outside the team can regulate—people outside the team cannot regulate" (external control). Participants' responses were recorded on a 1-to-9 bipolar scale with higher values representing attributions that were more internal, stable, team controllable, and externally controllable. In the current study, coefficient alpha reliabilities for the CDS-T

Table 3
Means/sums, standard deviations, coefficient alpha reliabilities, intra-class correlations, and bivariate correlations of attribution dimensions.

TRAMS	Team defeat			Team victory			TRAMS				CDS-T			
	M ± SD	α	ρ	M ± SD	α	ρ	C	S	G	U	LoC	Stab	TC	EC
C	3.67 ± .91	.83	.22*	3.76 ± .78	.78	.04		.26**	.43**	.11	.50**	.17*	.60**	−.28**
S	2.73 ± .89	.82	.22*	3.06 ± .99	.87	.13*	−.02		.31**	.02	.09	.62**	.12	.16*
G	3.41 ± .65	.65	.04	3.81 ± .67	.77	.13*	.16	.11		.44**	.46**	.14*	.40**	−.13
U	3.40 ± .73	.74	.02	3.65 ± .85	.84	.15*	.17	.18	.39**		.21**	−.12	.18*	<.01
CDS-T	SUM ± SD	α	ρ	SUM ± SD	α	ρ								
LoC	25.63 ± 6.98	.85	.30*	28.61 ± 5.37	.78	.12*	.37**	.11	.19*	.19*		.17*	.73**	−.26**
Stab	15.33 ± 6.41	.73	.28*	21.05 ± 6.69	.74	.19*	−.11	.52**	.11	.19*	.17		.11	.18**
TC	26.81 ± 7.44	.88	.29*	28.97 ± 6.09	.89	.03	.64**	.03	.19	.11	.54**	−.09		−.35**
EC	17.85 ± 7.02	.75	.21*	18.63 ± 7.11	.83	.04	−.08	.12	.02	−.03	−.13	.03	−.07	

Note. $n = 122$ for team defeat; $n = 212$ for team victory. C = controllability; S = stability; G = globality; U = universality; LoC = locus of causality; Stab = stability; TC = team control; EC = external control. α = coefficient alpha. ρ = intra-class correlation. Correlations (subscales were group-mean centred within teams) following team defeat are in the lower part of the correlation matrix and correlations following team victory are in the upper part of the correlation matrix.
$^{*}p < .05$. $^{**}p < .01$.

ranged from .72 to .88 and for the TRAMS ranged from .67 to .88 (see Table 3).

Analyses

The factor structure of the TRAMS was tested by analysing the pooled within-cluster covariance matrix, controlling for the nested nature of the data (Hox & Maas, 2001; Muthén, 1989).[8] For model fit, we examined measures of fit reported in Study 1. Correlations were used to determine relationships between measures. An alpha level of .05 was used for all tests.

Results

The sample comprised 191 participants from 28 losing teams, 241 participants from 42 winning teams, and 83 participants from 14 teams that tied. Sixty-nine participants from losing teams reported that they considered the competition a success and 29 participants from winning teams reported that they considered the competition a failure. These participants, along with those participants from teams that tied, were removed from further analyses. This resulted in a final data sample of 122 participants from 21 losing teams (team defeat: 29 female, $M_{age} = 22.03 \pm 4.16$ years, $M_{experience} = 10.93 \pm 5.01$ years, 72.00% White ethnicity) and 212 participants from 37 winning teams (team victory: 38 female, $M_{age} = 21.79 \pm 4.18$ years, $M_{experience} = 8.54 \pm 5.41$ years, 75.80% White ethnicity). The 58 teams included in analyses were from the sports of soccer ($n = 33$), ultimate Frisbee ($n = 7$), rugby ($n = 5$), hockey ($n = 4$), netball ($n = 4$), basketball ($n = 2$), cricket ($n = 2$), and American football ($n = 1$).

The factor structure of the TRAMS was tested using MPlus 7.11 (Muthén & Muthén, 1998–2012), imposing the TYPE = COMPLEX command to control for the nested nature of the data and resulted in modelling the asymptotic within-teams covariance matrix. Six participants from the team defeat condition and nine participants from the team victory condition were removed from analyses through listwise deletion for missing values. This resulted in 116 participants across 21 teams following defeat, and 211 participants across 37 teams following victory. An adequate and satisfactory fit (given the small sample) was observed for the TRAMS following defeat and a very good fit was observed following victory. Following defeat, although the chi-square statistic was significant ($\chi^2(84) = 155.40$, $p < .01$) and the CFI low (.86), the RMSEA was satisfactory (.08) and the SRMR was adequate (.07); following victory, the chi-square statistic was non-significant ($\chi^2(84) = 93.70$, $p > .05$), the RMSEA was low (.02) with a non-significant test for close fit, the SRMR was low (.04), and the CFI (.99) was high.[9]

For the correlation analysis, missing values for TRAMS and CDS-T items were replaced using the expectation-maximization procedure. Means/Sums, standard deviations, coefficient alpha reliabilities, intraclass correlations, and correlations of attribution dimensions are reported in Table 3. To control for the non-independence of data points (individual data were nested within teams), TRAMS and CDS-T subscales were group-mean centred within teams prior to correlation analyses. Within the TRAMS, correlations between factors ranged from low to moderate ($r = .02–.44$); within the CDS-T, correlations between factors ranged from low to high ($r = .03–.73$). Across measures, correlations between factors ranged from $r = <.01$ to .64. The hypothesized relationships were significant and in the predicted direction: The controllability subscale of the TRAMS was significantly and positively associated with the team control subscale of the CDS-T following both team defeat ($r = .64$, $p < .01$) and team victory ($r = .60$, $p < .01$), and the stability subscale of the TRAMS was significantly and positively associated with the stability subscale of the CDS-T following both team defeat ($r = .52$, $p < .01$) and team victory ($r = .62$, $p < .01$). In summary, all four correlations that were hypothesized to be significant were supported with moderately strong correlations (r's = .52–.64). Further, none of the 28 remaining (non-hypothesized) correlations across the measures were greater than $r = .52$ (i.e., the weakest hypothesized correlation). The proportion (100%) of correlations below the weakest hypothesized correlation ($r = .52$) was significantly different than might be expected due to chance.

Discussion

The results provided evidence to support the study hypotheses: The controllability and stability subscales of the TRAMS were significantly and positively associated with the team control and stability subscales of the CDS-T, respectively. Collectively, the

[8] Although multilevel confirmatory factor analyses, where the within-group and the between-group variance is modelled simultaneously, is regarded as the most appropriate method to examine the factor structure of measurement models where the data are meaningfully nested (e.g., Muthén, 1989), it requires large level two samples. Indeed Hox and Maas (2001) suggested that the level two sample size should be $N \geq 100$ (i.e., 100 + teams).

[9] The fit indices for the CDS-T following defeat were $\chi^2(98) = 128.41$, $p < .05$; RMSEA = .05, $p > .05$; SRMR = .08; CFI = .96; and, following victory were $\chi^2(98) = 233.47$, $p < .01$; RMSEA = .08, $p < .01$; SRMR = .08; CFI = .88.

results provided further evidence of construct validity for the TRAMS: Further support was provided for the factor structure of the TRAMS with independent samples across both team defeat (tentative support was provided with a small sample) and team victory conditions, together with initial evidence of concurrent validity for the controllability and stability subscales of the TRAMS.

Study 3

Method

Participants

Participants were 165 athletes (25 female) from 19 teams ($M_{age} = 26.39 \pm 11.35$ years). The sample was predominantly White (73.33%) and had an average length of experience in their sport for 12.92 years (SD = 8.82 years). Participants competed at club ($n = 100$), county ($n = 33$), regional ($n = 19$), national ($n = 6$), or international ($n = 1$) level (six participants did not report their level of performance). The 19 teams sampled were from the sports of cricket ($n = 9$), soccer ($n = 5$), rugby ($n = 3$), basketball ($n = 1$), and bowls ($n = 1$).

Procedure

Ethical approval for the study was awarded by a university research ethics board prior to data collection. The majority of competitions were league games. For each team, we collected data at three time points. At Time 1 (Day 1), one hour before performance (to allow participants time to prepare for the competition), participants completed a measure of collective-efficacy relating to an up-coming competition. This was regarded as participants' pre-competition collective-efficacy. At Time 2 (Day 1), one hour after performance (to give participants a chance to physically recover from competition), participants were asked, “Would you consider your team's performance a success or failure?” with binary responses of “success” and “failure”. Participants were also asked, “In regard to the outcome of the game or competition, did you win, lose or draw?” with response options of “win”, “lose” and “draw”. An open-ended statement required participants to write down the single most important reason to explain their team performance. In relation to this reason, participants completed a measure of attributions. At Time 3 (Day 7–9), one hour before performance (to allow participants time to prepare for the competition), participants completed a measure of collective-efficacy relating to an up-coming competition (performances at Days 1 and 7–9 were successive). This was regarded as participants' subsequent collective-efficacy.

Measures

Attributions. Team-referent attributions were assessed using the 15-item four-factor TRAMS developed and confirmed in Studies 1 and 2.

Collective-efficacy. Collective-efficacy was assessed using the Collective Efficacy Questionnaire for Sport (CEQS; Short, Sullivan, & Feltz, 2005). The CEQS comprises 20 items assessing five subcomponents of collective-efficacy: ability (e.g. “outplay the opposing team”), effort (e.g. “demonstrate a strong work ethic”), persistence (e.g. “perform under pressure”), preparation (e.g. “mentally prepare for this competition”) and unity (e.g. “keep a positive attitude”). The subscales can also be combined to create a composite collective-efficacy score; in the present study we used the composite score only. Responses were recorded on a ten-point bipolar scale anchored by the word-pairing: “*not at all confident*” to “*extremely confident*”. Short et al. reported an acceptable fit for the factor structure of the CEQS with adult sport performers: $\chi^2(160) = 574.29$, NNFI = .90, CFI = .92, SRMR = .04, RMSEA = .09.

Analyses

The sample was split into 83 participants from losing teams (10 teams) and 82 participants from winning teams (nine teams). All participants in winning teams considered the competition a success and 79 of the 83 participants in losing teams considered the competition a failure (the four participants that considered their team's defeat a success were removed from analyses). Eighteen participants from losing teams and 12 participants from winning teams were removed from the sample through listwise deletion for missing values (where participants had not competed in both competitions; remaining missing values for items were replaced using the expectation-maximization procedure). This resulted in a final data sample of 61 participants from ten losing teams (team defeat: all male, $M_{age} = 26.92 \pm 10.24$ years, $M_{experience} = 14.55 \pm 9.92$ years, 93.40% White ethnicity) and 70 participants from nine winning teams (team victory: 25 female, $M_{age} = 26.29 \pm 12.98$ years, $M_{experience} = 11.28 \pm 7.46$ years, 68.57% White ethnicity).

We controlled for the nested, interdependent data through running two-level regression models (variance estimates separated within-teams and between-teams), with the purpose of exploring associations between team-referent attribution dimensions (predictor variables) and subsequent collective-efficacy (criterion variable). The data were analysed using MLwiN 2.29 (Rasbash, Browne, Healy, Cameron, & Charlton, 2013) and estimates were calculated using the Restricted Iterative Generalised Least Squares (RIGLS) algorithm.[10] Using a similar design to self-referent attribution research (e.g., Coffee & Rees, 2008a) we entered our predictor variables in sequential steps. First, we controlled for pre-competition collective-efficacy by entering it independently at Step 1. Attribution dimension main effects were added at Step 2 and two-way interactive effects of controllability and generalisability dimensions were added at Step 3. Predictor variables were grand-mean centred and we used the change in the log-likelihood estimate and individual beta weights (and their standard errors) to ascertain significance. Regions of significance were computed for interactions using procedures specified by Preacher, Curran, and Bauer (2006). Prior to each analysis, data were checked for normality and homoscedasticity through visual inspection of standardised residual plots (against normal scores and fixed part predictions). In each case, data appeared normal and homoscedastic with no obvious outliers.

Results

Individual-level means, standard deviations, reliability coefficients, intra-class correlations, and bivariate correlations are reported in Table 4. Table 5 presents the results of the multilevel analyses. In the team defeat condition, pre-competition collective-efficacy was related to subsequent collective-efficacy, explaining 47.86% of the total residual variance ($b = .90$, $s_{\bar{x}} = .11$, $p < .01$). No significant improvement in model fit was shown at Step 2, $\Delta\chi^2(4) = 1.63$, $p > .05$, and no significant regression coefficients were observed. A significant improvement in model fit was observed, however, at Step 3, $\Delta\chi^2(3) = 9.90$, $p < .05$, with the interaction of controllability and globality ($b = -.47$, $s_{\bar{x}} = .21$,

[10] RIGLS is a modification to the standard IGLS (Iterative Generalised Least Squares) algorithm. In small samples, IGLS can produce biased estimates of the variance parameters and in such cases the RIGLS algorithm is preferred (Rasbash, Steele, Browne, & Goldstein, 2012).

Table 4
Individual-level descriptive statistics, reliability coefficients, intra-class correlations, and bivariate correlations.

	Team defeat			Team victory			Pre-CE	Sub-CE	C	S	G	U
	$M \pm$ SD	α	ρ	$M \pm$ SD	α	ρ						
Pre-CE	7.41 ± 1.16	.94	.17	7.69 ± .99	.94	.11		.68**	.38**	.13	.20	.26*
Sub-CE	7.04 ± 1.45	.97	.15	8.03 ± .85	.95	.14	.70**		.28*	.28*	.19	.25*
C	3.56 ± .76	.69	<.01	3.72 ± .72	.80	.03	.62**	.37**		.19	.36**	.45**
S	2.96 ± .68	.63	.06	3.31 ± .77	.76	.14	.40**	.42**	.40**		.53**	.37**
G	3.53 ± .78	.80	<.01	3.75 ± .58	.74	.11	.43**	.29*	.66**	.34**		.69**
U	3.43 ± .73	.76	<.01	3.67 ± .66	.83	.09	.46**	.37**	.48**	.41**	.66**	

Note. $n = 61$ for team defeat; $n = 70$ for team victory. Pre-CE = precompetition collective-efficacy, Sub-CE = subsequent collective-efficacy, C = controllability, S = stability, G = globality, U = universality. α = coefficient alpha. ρ = intra-class correlation. Correlations following team defeat are in the lower part of the correlation matrix and correlations following team victory are in the upper part of the correlation matrix.
$^*p < .05$. $^{**}p < .01$.

$p < .05$) and the interaction of controllability and universality ($b = .43$, $s_{\bar{x}} = .17$, $p < .01$) identified as salient predictors (explaining 3.45% additional collective variance). These interactions are depicted in Fig. 1a and b, respectively, using standard convention (values of −1 SD below the mean and +1 SD above the mean to indicate low and high levels of variables, respectively). Both figures demonstrate that controllability had a significant effect upon subsequent collective efficacy when causes generalised across situations and/or teams. Simple slopes were significant below −4.83 and above .70 SDs in levels of globality, and below −3.62 and above .65 SDs in levels of universality.

In the team victory condition, precompetition collective-efficacy explained 46.26% of the total residual variance in subsequent collective-efficacy ($b = .57$, $s_{\bar{x}} = .08$, $p < .01$). The addition of attribution dimension main effects at Step 2, $\Delta\chi^2(4) = 4.59$, $p > .05$, and interactive effects at Step 3, $\Delta\chi^2(3) = .63$, $p > .05$, did not significantly improve the overall model fit, but a significant positive regression coefficient was evident for stability at Step 2 ($b = .24$, $s_{\bar{x}} = .12$, $p < .05$).

Discussion

The results provided preliminary evidence to support the predictive validity of the TRAMS. Following team defeat, interactive effects for controllability and generalisability dimensions demonstrated that higher levels of controllability were associated with higher levels of subsequent collective-efficacy when causes were perceived to affect a narrow range of situations or to generalise across teams. Following team victory, a stable attribution was associated with higher levels of subsequent collective-efficacy. These findings are similar to those demonstrated for self-referent attributions and self-efficacy (Coffee & Rees, 2008a, 2008b, 2009) and support propositions that team-referent attributions are important for collective-efficacy in team sport (e.g., Greenlees et al., 2005).

Table 5
Multilevel regression models reporting the contribution of pre-competition collective-efficacy, and attribution dimensions to subsequent collective-efficacy following team defeat and team victory.

	Team defeat			Team victory		
	−2*log (χ^2)	$\Delta\chi^2$	b ($s_{\bar{x}}$)	−2*log (χ^2)	$\Delta\chi^2$	b ($s_{\bar{x}}$)
Intercept (random)	214.25		7.03 (.26)**	171.44		8.03 (.15)**
Step 1	170.40	43.85**		129.82	41.62**	
Pre-CE			.90 (.11)**			.57 (.08)**
Step 2	168.77	1.63		125.23	4.59	
C			−.22 (.26)			−.01 (.12)
S			.24 (.22)			.24 (.12)*
G			.06 (.20)			−.15 (.20)
U			.01 (.26)			.09 (.16)
Step 3	158.87	9.90*		124.60	.63	
CS			.21 (.15)			.06 (.10)
CG			−.47 (.21)*			−.01 (.12)
CU			.43 (.17)**			.00 (.11)

Note. $ni = 70$ and $nj = 9$ for team victory; $ni = 61$ and $nj = 10$ for team defeat. Pre-CE = pre-competition collective-efficacy, C = controllability, S = stability, G = globality, U = universality, CS = controllability*stability, CG = controllability*globality, CU = controllability*universality. The models presented are random intercept-fixed slopes models. Random slopes were also explored but did not significantly improve the model fit (no significant covariance's between intercepts and slopes).
$^*p < .05$. $^{**}p < .01$.

General discussion

This article presented three studies that together provide preliminary evidence for the construct, concurrent, and predictive validity of a new four-factor measure of team-referent attributions. Study 1 reported good fit indices for the TRAMS across least

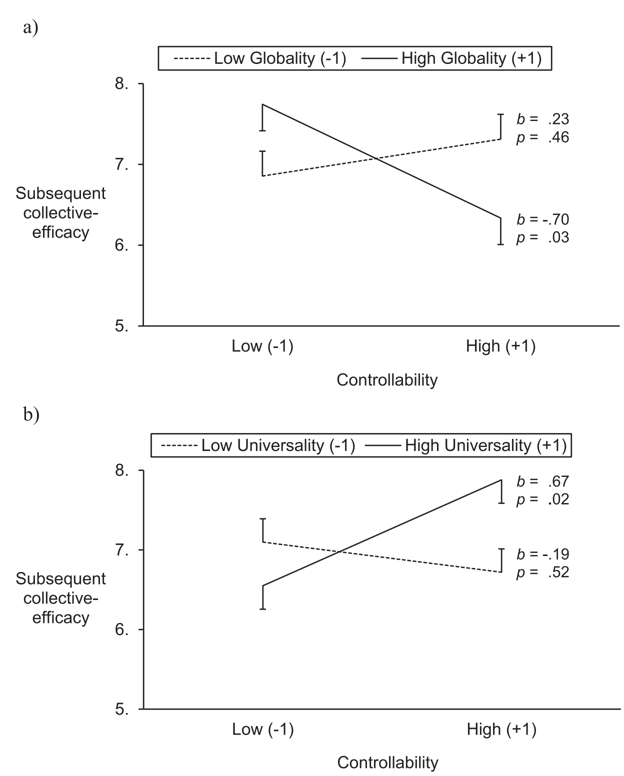

Fig. 1. The interactive effects for controllability and globality, and controllability and universality upon subsequent collective-efficacy (controlling for all other variables entered at steps 1, 2 and 3).

successful and most successful conditions. Study 2 also reported, with independent samples, evidence for the factor structure of the TRAMS across team defeat and team victory conditions, together with significant hypothesized positive correlations between TRAMS subscales and corresponding subscales of the CDS-T. The potential theoretical and applied advancements for examining an expanded conceptualisation of generalisability within team environments, together with examining interactive effects of attribution dimensions were addressed in Study 3.

Study 3 reported, following team defeat, interactive effects for controllability and generalisability dimensions upon subsequent collective-efficacy; following team victory, a significant positive coefficient was observed for the generalisability dimension of stability upon subsequent collective-efficacy. The interactions for controllability and the generalisability dimensions demonstrated that controllability had a significant effect upon subsequent collective-efficacy when causes of team defeat were also perceived to generalise across situations and/or across teams. Moreover, higher levels of controllability were associated with higher levels of subsequent collective-efficacy when causes were perceived to affect a narrow range of situations or were perceived as common to all teams (i.e., occur generally across teams). In other words, when team members perceived that the causes of team defeat were specific to the situation and/or were common causes of defeat across teams, high perceptions of controllability were associated with higher levels of subsequent collective-efficacy.

These results build upon research using the CDS-T that report interactive effects for external control and stability, and team control and stability upon subsequent collective-efficacy (Allen et al., 2009). The interactions reported by Allen and colleagues demonstrated that, when causes were perceived to generalise across time (were more stable), higher levels of control (team—following team defeat—and external—following team victory) were associated with higher levels of subsequent collective-efficacy. The current paper provides the first empirical attempts in team-referent attribution research to test for interactive effects of controllability and globality, and controllability and universality attributions. Although, in research on self-referent attributions, empirical evidence of interactions for controllability and globality have been reported (e.g., Coffee & Rees, 2008b), no evidence has been provided for interactive effects of controllability and universality. In comparison to the self-referent attribution literature, the current paper suggests that, for team-referent attributions, universality could be an important attribution dimension when exploring the generalisability of controllability attributions.

Following team victory, a significant coefficient was observed for stability upon subsequent collective-efficacy. In other words, following team victory, team members reported higher levels of collective-efficacy when they perceived causes of team performance as unlikely to change (or likely to recur). These results concur with Dithurbide et al. (2009) who, using the CDS-T, reported an interaction of objective performance and stability upon subsequent collective-efficacy, such that, only following successful performance, a stable attribution was positively related to subsequent collective-efficacy. Further, the results are similar to those demonstrated for the effects of self-referent attributions upon self-efficacy. For example, Coffee and Rees (2008b) demonstrated that, following more successful performances, higher levels of subsequent self-efficacy were more likely if causes for more successful performances were considered to generalise, either across time (stable) or situations (global), or were considered unique to the individual (personal).

As noted already, we found interactive effects of attributions following team defeat and a significant coefficient for stability following team victory. This is in line with propositions (e.g., Allen et al., 2012) and empirical evidence (e.g., Coffee & Rees, 2008a, 2008b, 2009) that interactive effects of attribution dimensions are important following defeat and that main effects of generalisability dimensions are important following victory. To some extent, this is underpinned by evidence that suggests that the attribution process is more salient following less successful performances (see, Wong & Weiner, 1981). At this point, caution should be exercised against generalising this pattern of effects based upon Study 3 because, although data were collected through a multiple time point design, the sample was limited in size. Further, following team victory, although the regression coefficient for stability was significant, collectively, the four attribution dimension main effects did not significantly improve the overall model fit.

The TRAMS was developed in line with the methodology employed by Greenlees et al. (2005) in the development of the CDS-T. That is, the TRAMS was developed by rewording items from the CSGU (a measure of self-referent attributions in sport) to reflect team-rather than self-referent attributions. This approach benefits the future development of the attribution literature, enabling self-referent and team-referent attribution research to be compared and understood with the knowledge that differences found will be due solely to the attribution perspective (team-referent or self-referent) and not due to idiosyncrasies in the content of items across attribution perspectives. A more traditional approach to item generation (see, e.g., Eys, Loughead, Bray, & Carron, 2009) would have likely resulted in variances in item and, therefore, factor content across self-referent (the CSGU) and team-referent (the TRAMS) measures in the literature. Whilst our methodological approach to the development of the TRAMS permits congruent development of self- and team-referent attribution literature, it should be acknowledged that the methodology employed may not have identified all elements that are of relevance in the measurement of team-referent attributions.

The TRAMS is a measure of situation-specific team-referent attributions, distinguishable from measures of attributional style in which the purpose is to identify participants' cognitive predispositions to explaining the causes of events (Peterson et al., 1982). Within sport, Shapcott and Carron (2010) recently developed the Team Attributional Style Questionnaire (the TASQ). In line with the CSGU and the TRAMS, the TASQ draws upon Rees et al.'s (2005) four-dimensional attribution model (i.e., controllability, stability, globality, and universality) as its theoretical underpinning. The development of the TRAMS alongside the existence of the TASQ presents an opportunity for researchers to explore the effects of both situation-specific and dispositional tendencies of team-referent attributions.

In conclusion, the TRAMS is unique and at the same time complementary in its offering to advance attribution literature. The present study extended the applicability of Rees et al.'s (2005) conceptual model to team-referent, situation-specific attributions. Furthermore, the four-factor measure was found to be reliable and valid across independent samples, with evidence provided for the construct, concurrent, and predictive validity of the TRAMS. We hope that the development of the TRAMS will encourage researchers to further explore the main and interactive effects of team-referent attributions upon outcomes such as collective-efficacy, team cohesion, and social identity.

Acknowledgements

The authors thank Matt Slater, Sean Figgins, Sophie Martin, Mike Byrne, Adam Bond, Graeme Lavery, Conor Reid, and Justine Robinson for their help during data collection. We also thank Dr. Calum Arthur, University of Stirling, for his assistance with analysing data for Study 2.

Appendix A. Supplementary data

Supplementary data related to this article can be found at http://dx.doi.org/10.1016/j.psychsport.2014.10.009.

References

Allen, M. S., Coffee, P., & Greenlees, I. (2012). A theoretical framework and research agenda for studying team attributions in sport. *International Review of Sport and Exercise Psychology, 5*, 121–144. http://dx.doi.org/10.1080/1750984X.2012.663400.

Allen, M. S., Jones, M. V., & Sheffield, D. (2009). Attribution, emotion and collective-efficacy in sport teams. *Group Dynamics: Theory, Research, and Practice, 13*, 205–217. http://dx.doi.org/10.1037/a0015149.

Biddle, S. J. H., & Hanrahan, S. J. (1998). Attributions and attributional style. In J. L. Duda (Ed.), *Advances in sport and exercise psychology measurement* (pp. 3–19). Morgantown, WV: Fitness Information Technology.

Biddle, S. J. H., Hanrahan, S. J., & Sellars, C. N. (2001). Attributions: past, present, and future. In R. N. Singer, H. A. Hausenblas, & C. M. Janelle (Eds.), *Handbook of sport psychology* (2nd ed.). (pp. 444–471). New York: Wiley.

Browne, M. W., & Cudeck, R. (1993). Alternative ways of assessing model fit. In K. A. Bollen, & J. S. Long (Eds.), *Testing structural equation models* (pp. 136–162). Newbury Park, CA: Sage.

Byrne, B. M. (2006). *Structural equation modeling with EQS: Basic concepts, applications, and programming* (2nd ed.). Mahwah, NJ: Lawrence Erlbaum Associates.

Cheung, G. W., & Rensvold, R. B. (2002). Evaluating goodness-of-fit indexes for testing measurement invariance. *Structural Equation Modeling, 9*, 233–255. http://dx.doi.org/10.1207/S15328007SEM0902_5.

Chow, G. M., & Feltz, D. L. (2008). Exploring the relationships between collective efficacy, perceptions of success, and team attributions. *Journal of Sports Sciences, 26*, 1179–1189. http://dx.doi.org/10.1080/02640410802101827.

Coffee, P., & Rees, T. (2008a). The CSGU: a measure of controllability, stability, globality, and universality attributions. *Journal of Sport and Exercise Psychology, 30*, 611–641.

Coffee, P., & Rees, T. (2008b). Main and interactive effects of controllability and generalisability attributions upon self-efficacy. *Psychology of Sport and Exercise, 9*, 775–785. http://dx.doi.org/10.1016/j.psychsport.2007.12.002.

Coffee, P., & Rees, T. (2009). The main and interactive effects of immediate and reflective attributions upon subsequent self-efficacy. *European Journal of Sport Science, 9*, 41–52. http://dx.doi.org/10.1080/17461390802594227.

Crocker, P. R., Brune, S. M., Kowalski, K. C., Mack, D. E., Wilson, P. M., & Sabiston, C. M. (2014). Body-related state shame and guilt in women: do causal attributions mediate the influence of physical self-concept and shame and guilt proneness. *Body Image, 11*, 19–26. http://dx.doi.org/10.1016/j.bodyim.2013.08.002.

Crocker, P. R. E., Eklund, R. C., & Graham, T. R. (2002). Evaluating the factorial structure of the revised causal dimension scale in adolescents. *Research Quarterly for Exercise and Sport, 73*, 211–218. http://dx.doi.org/10.1080/02701367.2002.10609011.

Dithurbide, L., Sullivan, P., & Chow, G. (2009). Examining the influence of team referent causal attributions and team performance on collective-efficacy: a multilevel analysis. *Small Group Research, 40*, 491–507. http://dx.doi.org/10.1177/1046496409340328.

Eys, M., Loughead, T., Bray, S. R., & Carron, A. V. (2009). Development of a cohesion questionnaire for youth: the Youth Sport Environment Questionnaire. *Journal of Sport and Exercise Psychology, 31*, 390–408.

Fornell, C., & Larcker, D. G. (1981). Evaluating structural equation models with unobservable variables and measurement error. *Journal of Marketing Research, 18*, 39–50. http://dx.doi.org/10.2307/3151312.

Greenlees, I., Lane, A. M., Thelwell, R., Holder, T., & Hobson, G. (2005). Team-referent attributions among sport performers. *Research Quarterly for Exercise and Sport, 76*, 477–487.

Greenlees, I., Stopforth, M., Graydon, J., Thelwell, R., Filby, W., & El-Hakim, Y. (2007). The impact of match importance and gender on the team-serving attributional bias among interdependent sports team players. *Group Dynamics: Theory, Research, and Practice, 11*, 54–65. http://dx.doi.org/10.1037/1089-2699.11.1.54.

Hox, J. J., & Maas, C. J. M. (2001). The accuracy of multilevel structural equation modeling with pseudobalanced groups and small samples. *Structural Equation Modeling – A Multidisciplinary Journal, 8*, 157–174. http://dx.doi.org/10.1207/S15328007SEM0802_1.

Hu, L., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: conventional criteria versus new alternatives. *Structural Equation Modeling, 6*, 1–55. http://dx.doi.org/10.1080/10705519909540118.

Jöreskog, K. G. (1993). Testing structural equation models. In K. A. Bollen, & J. S. Long (Eds.), *Testing structural equation models* (pp. 294–316). Newbury Park, CA: Sage.

Jöreskog, K. G., & Sörbom, D. (1996). *LISREL 8 user's reference guide*. Chicago, IL: Scientific Software International.

Jöreskog, K. G., & Sörbom, D. (2012). *LISREL 9.10 for Windows [Computer software]*. Skokie, IL: Scientific Software International, Inc.

McAuley, E., Duncan, T., & Russell, D. (1992). Measuring causal attributions: the Revised Causal Dimension Scale (CDS-II). *Personality and Social Psychology Bulletin, 18*, 566–573. http://dx.doi.org/10.1177/0146167292185006.

Muthén, B. O. (1989). Latent variable modeling in heterogeneous populations. *Psychometrika, 54*, 557–585. http://dx.doi.org/10.1007/BF02296397.

Muthén, L. K., & Muthén, B. O. (1998-2012). *Mplus user's guide* (7th ed.). Los Angeles, CA: Muthén & Muthén.

Peterson, C., Semmel, A., von Baeyer, C., Abramson, L. T., Metalsky, G. I., & Seligman, M. E. P. (1982). The Attributional Style Questionnaire. *Cognitive Therapy and Research, 6*, 287–300.

Preacher, K. J., Curran, P. J., & Bauer, D. J. (2006). Computational tools for probing interaction effects in multiple linear regression, multilevel modeling, and latent curve analysis. *Journal of Educational and Behavioral Statistics, 31*, 437–448. http://dx.doi.org/10.3102/10769986031004437.

Rasbash, J., Browne, W. J., Healy, M., Cameron, B., & Charlton, C. (2013). *MLwiN Version 2.29 [Computer software]*. Centre for Multilevel Modelling, University of Bristol.

Rasbash, J., Steele, F., Browne, W. J., & Goldstein, H. (2012). *A user's guide to MLwiN: Version 2.26*. Bristol, England: Centre for Multilevel Modelling, University of Bristol.

Raykov, T. (2006). Examining temporal stability of scale validity in longitudinal studies. *Multivariate Behavioral Research, 41*, 401–426. http://dx.doi.org/10.1207/s15327906mbr4103_6.

Rees, T., Ingledew, D. K., & Hardy, L. (2005). Attribution in sport psychology: seeking congruence between theory, research and practice. *Psychology of Sport and Exercise, 6*, 189–204. http://dx.doi.org/10.1016/j.psychsport.2003.10.008.

Satorra, A., & Bentler, P. M. (2001). A scaled chi-square test statistic for moment structure analysis. *Psychometrika, 66*, 507–514. http://dx.doi.org/10.1007/BF02296192.

Shapcott, K. M., & Carron, A. V. (2010). Development and validation of a team attributional style questionnaire. *Group Dynamics: Theory, Research, and Practice, 14*, 93–113. http://dx.doi.org/10.1037/a0018252.

Shapcott, K. M., Carron, A. V., Greenlees, I., & El Hakim, Y. (2010). Determinants of team-referent attributions: a multi-level analysis. *International Journal of Sport and Exercise Psychology, 8*, 227–244. http://dx.doi.org/10.1080/1612197X.2010.9671951.

Shook, C. L., Ketchen, D. J., Hult, G. T. M., & Kacmar, K. M. (2004). An assessment of the use of structural equation modeling in strategic management research. *Strategic Management Journal, 25*, 397–404. http://dx.doi.org/10.1002/smj.385.

Short, S. E., Sullivan, P., & Feltz, D. L. (2005). Development and preliminary validation of the collective-efficacy questionnaire for sports. *Measurement in Physical Education and Exercise Science, 9*, 181–202. http://dx.doi.org/10.1207/s15327841mpee0903_3.

Weiner, B. (1985). An attributional theory of achievement motivation and emotion. *Psychological Review, 92*, 548–573. http://dx.doi.org/10.1037/0033-295X.92.4.548.

Wong, P. T. P., & Weiner, B. (1981). When people ask "why" questions, and the heuristics of attributional search. *Journal of Personality and Social Psychology, 40*, 650–663. http://dx.doi.org/10.1037/0022-3514.40.4.650.